7) AGE, an attempt to build: a software package of AI techniques and methods for problem solving and hypothesis formation; and its associated user-interface.

The superb computing facilities of the NIH-supported SUMEX-AIM timesharing facility will be available at no charge to this project. The SUMEX-AIM facility, with Prof. Lederberg as principal investigator, is a national resource for the application of artificial intelligence techniques to problems in biology and medicine. Resources to be provided will include all CPU-time and storage required. Those involved at Stanford will be operating through hard-wired or dial-up equipment to the SUMEX PDP-10, while those at the University of New Mexico will access the system through either the ARPA network or TYMNET.

The SUMEX-AIM facility is a powerful interactive computing system open to a national community. Interlisp and other high level languages are available and supported by a large system staff. Many convenient text editors for developing programs are provided. The TENEX operating system supports flexible file handling and sophisticated storage management for a highly interactive computing environment.

Appendix I

GLOSSARY

AGE — Stanford project ("Attempt to GEneralize") to build a general package of AI methods.

AI — Artificial Intelligence

AI "toolbox" — A set of AI concepts embodied in programs that are general enough to be used to construct problem solving programs in many different domains.

alkaline phosphatase — An enzyme that removes terminal phosphates from nucleic acids, and whose optimum pH is alkaline.

AT-rich DNA — DNA containing a high proportion of Adenine + Thymine base pairs. Because of the Watson-Crick base pairing rules, A/T=1 and G/C=1, but (A+T)/(G+C) ratios can can vary widely between DNA of different species and even different regions of the same DNA molecule.

B. subtilis — A common soil bacterium often used in genetic experiments.

bacteriophage — Viruses that multiply in bacteria.

base sequence — A string of nucleotides in a nucleic acid.

bottom-up process — A program that puts together inferences from data without the benefit of global expectations and goals.

CONGEN — Constrained Generator of molecular structures for DENDRAL.

dalton — A unit of mass equal to that of a single hydrogen atom.

data-driven procedure — bottom-up process

demon — a procedure in a program that is triggered by an event, as opposed to being executed in the "normal" execution of a sequential program.

denaturation — The loss of the native configuration of a macromolecule resulting, for example, from heat treatment, extreme pH changes, chemical treatment, or other denaturing agents. It is usually accompanied by the separation of strands (in DNA) and the loss of biological activity.

DENDRAL — Heuristic program for generating and testing organic molecular structures/as candidate explanations of empirical data.

digestion — With reference to enzymes, implies the cleaving of chemical bonds in the target molecule. For example, exonucleases "erode" or remove terminal nucleotides, restriction enzymes cut at the internal recognition sequences.

dimer — A concatenated DNA structure consisting of two identical constituents.

discrimination experiment — A series of experimental steps designed to conclude whether structures are identical or not.

DNA — Deoxyribonucleic acid. A polymer of deoxyribonucleotides (see nucleotide definition). Can exist as double or single strands. The genetic material of all cells and the central molecule in molecular genetics.

domain-specific critic — A procedure which applies specific genetics knowledge to problem solving, as opposed to general problem-solving knowledge.

EcoRl — A restriction enzyme isolated from a strain of E. coli that cleaves DNA at site-specific regions along the molecule. Its recognition site is 5'-GAATTC-3'.

EDNA — The DNA-structure-editor for MOLGEN.

electron microscopy — Abbreviated EM. A high-resolution technique for visualizing material that uses beams of electrons instead of light rays. Resolutions of about 10^-7 cm are posible with biological materials.

electrophoresis — An experimental technique used to separate, purify, and measure the molecular weight of molecules having an electric charge in solution.

endonuclease — An enzyme that cuts DNA backbone chains internally.

enzyme — Protein molecule capable of catalyzing a specific chemical reaction.

E. coli — A common intestinal bacterium: the most intensively studied organism except for man.

event-driven procedure — demon

exonuclease — An enzyme that digests DNA from the ends of strands.

experiment planning — An activity characterized by the production of a sequence of experimental steps to achieve a goal.

focus rules — Focus of attention procedures. Items of knowledge that guide a program to the most relevant parts of the problem or the most useful subroutines.

gaps — An internal feature of double-stranded DNA which is a region of unpaired nucleotides due to the excision of a string on one strand.

HEARSAY — AI program written at Carnegie-Mellon University to understand spoken English. Integrates inferences made by multiple experts.

hierarchical planning — AI techniques refined by Sacerdoti which uses a hierarchy of descriptions to plan an efficient problem solution procedure.

inspector — Domain-specific critic

INTERLISP — A powerful extension of the LISP programming language.

KRL-0 — A programming language (knowledge representation language) developed at Stanford and Xerox, Palo Alto Research Center.

ligase — An enzyme capable of covalently joining parts of, or entire DNA molecules together.

ligation — The enzymatic joining together of DNA molecules.

linear DNA — Double stranded DNA that is not covalently closed at its termini.

meta-rules — Rules for a program that mentions domain-specific rules, i.e, to prune or reorder the set of rules relevant for problem solving in specific contexts.

molecular adapter — A chemically synthesized segment of DNA that is utilized to join together DNA molecules which do not have complementary termini for ligation.

MOLGEN — Computer program for reasoning in molecular genetics. Main subject of this proposal.

monomer — A single DNA molecule (or nucleotide) that has not undergone polymerization (viz. a unit character capable of assembly into a string).

MYCIN — Medical diagnosis and therapy recommendation program developed at Stanford.

nicks — A local interruption in the phosphodiester backbone of DNA. No genetic information is missing due to this structural anomaly.

nuclease — An enzyme which breaks chemical bonds in the DNA phosphodiester backbone. Consists of endonucleases and exonucleases.

nucleotide — The building blocks of DNA consisting of a purine (Adenine or Guanine) or a pyrimidine (Thymine or Cytosine) linked to a deoxyribose sugar with a phosphate group also linked to adjacent sugar. Adjacent nucleotides are linked together through a phosphate group and a hydroxyl group on the sugar component (see phosphodiester).

pH — The negative logarithm of the effective hydrogen ion concentration or hydrogen ion activity in gram equivalents per liter. Used in expressing both acidity and alkalinity on a scale whose values run from 0 to 14; 7 representing neutrality, less than 7 increasing acidity, >7 increasing alkalinity. DNA exists in native form between pH values of 5 and 12.

phosphodiester — The chemical link between adjacent nucleotides. The following diagram of its structure was drawn using CONGEN [6]:

```
                                .
                               .
          .                   .
           .                 /
            .           O   O
             \           \ /
              O   O       P
               \ /       / \\
                P       O   O
               // \    /
              O   O  CH2
                   \ /
                   C-C
                  /  |
                CH   |
                  \  |
                   C-C
                  /
                BASE
```

plan schema — A sketch of a procedure describing the plan for an experiment in abstract, general terms.

planning islands — Partial solutions to a problem found by a planning program. "Stepping stones" to a complete solution.

planning rules — Procedures or items of knowledge that aid a program in constructing a problem solving plan.

plasmid — Extrachromosomal DNA molecules which are double stranded, circular, and supercoiled. They range in size from about 5*10^6 daltons to near 10^8. Small plasmids can exist in many (more than 50) copies per cell while large ones are maintained at one or two. They are often used as vectors for amplifying and transferring DNA from one organism to another.

poly-A region (sequence) — A homopolymeric sequence of adenine nucleotides. Implies a poly-T region on the complementary strand.

poly-C — Homopolymeric cytidine nucleotides.

poly-G — Homopolymeric guanine mucleotides.

poly-T — Homopolymeric thymine nucleoties.

polymerase — Enzymes that are catalysts for nucleic acid chain growth.

pre-conditions — Premise clauses of conditional sentences that must be satisfied before the consequent actions are taken.

production rules — Conditional sentences used to encode inferential knowledge for a program.

prototype — The type of unit created for representing information about general concepts. Features are defined by slots associated wih the prototype.

restriction enzyme — Site-specific endonucleases used frequently in molecular genetic manipulations. Allow previously impossible experiments to be performed due to their ability to cleave DNA at reproducible locations allowing rearrangements within and between molecules.

RNA — Ribonucleic acid. Typically single stranded, is a polymer of ribonucleotides connected by phosphodiester bonds.

schema/rule schema/program schema — An abstract, generalized representation of a concept or program. In MOLGEN, program schemata (or rule schemata) are represented as Units with slots defined for important features.

SECS — Chemical synthesis planning program developed by Prof. Todd Wipke (U.C. Santa Cruz).

self-circularization — Ligation of the ends of the same DNA resulting in a circular, covalently closed molecule.

self-ligation — Ligation of a DNA molecule to itself, resulting in in a circular molecule. Catalyzed by ligase.

sequencing experiment — A technique to determine the order of nucleotides in a strand of DNA.

slots — Pre-defined features of objects for which values are sought.

SMALLTALK — Display-oriented programming language developed at Xerox, Palo Alto Research Center.

sticky ends	A condition of partial single-strandedness at the termini of DNA molecules, allowing base pairing in that region. Restriction enzymes often leave sticky ends, greatly facilitating the rearrangements of DNA.
STRIPS	Robot plnning program developed at SRI.
SUMEX-AIM	NIH-sponsored computer resource for applications of artificial intelligence in medicine.
Teiresias	AI program that acquires inference rules for MYCIN and guides MYCIN reasoning.
TENEX	Operating system for the DEC KI-10 system running at the SUMEX-AIM facility.
Thy gene	A gene coding for the enzyme, thymidylate synthetase. This enzyme is crucial in enabling a bacterium possessing it to produce thymidine, a constituent of DNA.
top-down process	A program that works from general principles, testing data against expectations and goals, often working by dividing complex problems into simpler ones.
Units	Basic element of representation in MOLGEN. Units are organized in a hierarchy to facilitate the representation of class-subclass and prototype-instance relationships. Units are used for representing processes as well as concepts.
vector	A self-propagating DNA molecule that can be used to link DNA sequences of interest. Vectors can be one of several replicating plasmid or bacteriophage DNAs.
world states	Representation of the state of an experiment at any given time. The "world" for the program is the limited set of objects and operations relevant to a specific experiment.
3'-end/ 5'-end	Related to the direction of the phosphodiester bonds in the backbone of DNA molecules. Each strand thus has one 3' end and one 5' end.

Appendix II

EDNA -- The Editor for DNA

The following is an actual session with the MOLGEN knowledge acquisition system recorded from SUMEX. Comments preceded by a semicolon have been inserted to clarify some aspects of the dialog.

```
@ue                                        ;UE is the name of the Unit
                                           ;Editor. Here it is being called
(Version  4-OCT-77 08:56:16)               ;from TENEX

Welcome to the MOLGEN Unit Editor.  Type ? anytime for assistance.
The symbol : indicates that the editor is waiting for your input.
Two characters are enough for command recognition.  You may type ahead
responses for a command.

Name of Network: jerry                      ;Jerry is the name of an
                                            ;existing Knowledge Base on
                                            ;file.
:create testl root specialization           ;A new unit TEST1 is created.
Give a value for the DESCR slot             ;UE asks for documentation.
Text Editor
te: Test unit to demonstrate the DNA structure editor
te: done                                    ;User indicates he is done
                                            ;with documentation.
Do you want to see what slots have already been filled? yes

DESCR:           (U)    from ROOT     <DESCR>
Test unit to demonstrate the DNA structure editor
MODIFIER:        (U)    from ROOT     <MODIFIER>  STEFIK
CREATOR:         (U)    from ROOT     <CREATOR>   STEFIK
MODIFIED:        (U)    from ROOT     <MODIFIED>   6-OCT-77 10:22:03
CREATED:         (U)    from ROOT     <CREATED>    6-OCT-77 10:22:03

                                            ;Note that the system has
                                            ;automatically recorded the
                                            ;author, date and time of the
                                            ;new unit.

You can now create new slots or edit old ones.  When through type DONE
EDIT: create substrate
Datatype: dna                               ;DNA is a datatype.
Role: r                                     ;"Role" controls transmission
                                            ;of the value in the
                                            ;substrate slot if we make
```

```
                                        ;specializations of TEST1.
Is it a dynamic slot? n
DNA Editor                              ;Since the datatype is DNA,
                                        ;we get the DNA editor.
Copy or Create anew? create
Segment Type: ?                         ;A "?" may always be typed to
                                        ;tell the system to clarify
                                        ;what it expects for a
                                        ;response.

Choose one of the following Segment Types.

Type    Description
LE      Length Segment.  Indicates number of nucleotides in a region.
BA      Base Segment.  Indicates actual Base Sequence.
SI      Cut Site for enzyme.
Segment Type: length
Length: ?                               ;Another "?"

Indicate number of nucleotides as in the following examples:

You Type        Meaning

5               5 nucleotides
100 200         Between 100 and 200 nucleotides.
1K  1.3K        Between 1000 and 1300 nucleotides. (K=1000) No Spaces!
R 1K 1.3K       Same as above except RNA instead of DNA.
Length: 2.5k
                                        ;The initial segment is now
                                        ;specified.  We may now issue
                                        ;any legal EDNA command.
edna: print

DNA Printer (Version 19-SEP-77 )
     1
     (2500)
     ------

edna: insert 1 3' bases attacg
edna: print

     1      2 3 4 5 6 7                 ;EDNA presents its structures
     (2500) A T T A C G                 ;pictorially.  Segments are
     ------------------                 ;referenced by number.

edna: mirror 4 to 7                     ;"Mirror" means to add a
edna: print                             ;parallel DNA strand.

     1      2 3 4  5  6 7
     (2500) A T T  A  C G               ;EDNA knows about
```

```
     --------------------              ;complementary bases.
               ---------
               A  T  G C
               11 10 9 8

edna: break 4 h                        ;Break a bond.  A Break
edna: print                            ;command can specify 5', 3',
                                       ;or H bonds.
     1      2 3 4 5  6 7
     (2500) A T T A  C G
     ------------------
                /------                ;Segment 11 is depressed
              -- T  G C                ;to indicate the broken bond.
              A  10 9 8
              11

edna: connect 7 3' 8
edna: print

     1      2 3 4 5  6 7
     (2500) A T T A  C G->8            ;A similar notation is used
     --------------------|             ;to indicate Hairpin loops.
                /--------|
              -- T  G C
              A  10 9 8
              11

edna: undo                             ;EDNA can "undo" any of its
                                       ;structure changing commands
CONNECT undone.
edna: done                             ;User is finished editing
                                       ;this structure.  He returns to
                                       ;the slot editor.
EDIT: print all
DESCR:            (U)   from ROOT     <DESCR>
Test unit to demonstrate the DNA structure editor
MODIFIER:         (U)   from ROOT     <MODIFIER>  STEFIK
CREATOR:          (U)   from ROOT     <CREATOR>   STEFIK
MODIFIED:         (U)   from ROOT     <MODIFIED>   6-OCT-77 10:22:09
CREATED:          (U)   from ROOT     <CREATED>    6-OCT-77 10:22:03
SUBSTRATE:        (R)   *Top*         <DNA>       (Renumbering 11)

     1      2 3 4 5  6 7
     (2500) A T T A  C G
     ------------------
                /------
              -- T  G C
              A  10 9 8
              11
```

```
EDIT: done                    ;User is done with this unit.
:done                         ;User is done with this
                              ;knowledge base.
Save JERRY? no                ;He doesn't save his changes
Bye                           ;because this was just a
                              ;demonstration.
```

Appendix III

A Genetic Planning Example

This section is intended to extend the range of genetic examples for which MOLGEN is envisioned being applied. In particular, the recent cloning of the rat insulin gene in E. coli [38] has been achieved using a simple additional step to the usual experimental protocol. It is asserted that the genesis of this efficiency-improving step can be found in the relatively simple application of knowledge about enzyme properties and DNA ligation kinetics.

The basic experimental outline is seen below, modified from [38]. It closely follows the 'classical' recombinant DNA methodology, with a few additional steps. The one we wish to focus on most closely in this discussion is the application of Alkaline Phosphatase to the plasmid vector after cutting of the plasmid by the restriction enzyme HindIII.

```
  mRNA
~~~~~~~~(A)x 3'

      |
      | Reverse transcriptase
      |

- ~~~~~~~(A)x 3'
|
---------(T)y 5'

       |
       | Alkaline Digestion
       |

- 3'
|
---------(T)y 5'

        |
        | Reverse transcriptase
        |
```

```
---------(A)z 3'
|
---------(T)y 5'                                        plasmid

          |
          | S1 Nuclease                                  HindIII site
          |                                                --//--
                                                        / -//--- \
---------(A)z 3'       cDNA                             ||      ||
                                                        ||      ||
---------(T)z 5'                                        ||      ||
                                                        ||      ||
          |                                             \ ------ /
          | T4 DNA Ligase                                 ------
          |      +
          | 5' CCAAGCTTGG 3'                              | HindIII
          | 3' GGTTCGAACC 5'                              | restriction
          |                                               | enzyme

5' CCAAGCTTGG---------(A)zCCAAGCTTGG 3'          5' pAGCTT-----A  3'
3' GGTTCGAACC---------(T)zGGTTCGAACC 5'              3' A-----TTCGAp 5'

          |                                               |
          | HindIII restriction enzyme                    | ALKALINE
          |                                               | PHOSPHATASE

5' pAGCTTGG---------(A)zCCA 3'                  5' ohAGCTT-----A 3'
     3' ACC---------(T)zGGTTCGAp 5'                 3' A-----TTCGAoh 5'

          |                                               |
           \                                              |
            \                                            /
             \                                          /
               - - - - - - - - - - - - - - - -   - - -
                                              \  /
                                               \/
                                               | T4 DNA Ligase
                                               |
                                                  *
                  --AAGCTTGG---------(A)zCCAAGCTT--
                / --TTCGAACC---------(A)zGGTTCGAA-- \
                ||       *      cDNA                 ||
                ||                                   ||
                ||                                   ||
                ||          plasmid DNA              ||
                \ ---------------------------------- /
                  ----------------------------------
```

The following steps were carried out. First, insulin messenger RNA was purified from B cells in the rat pancreas. This was reverse transcribed into a hybrid DNA/RNA structure by the use of avian myeloblastosis virus (AMV) reverse transcriptase and the RNA selectively degraded by raising the pH. A double-stranded DNA form was synthesized by incubating this with deoxynucleoside triphosphates and the AMV reverse transcriptase (a DNA polymerase could have been used). The hairpin at the end of the molecules and any non-base paired regions were removed with the single-strand specific nuclease S1.

The resulting molecular structure is termed cDNA, or copy DNA, because it should contain the precise genetic information contained in the gene coding for the insulin messenger RNA. This is the in-vitro synthesized segment that is to be cloned in bacterial recipients for amplification and analysis.

A recently developed technique for ligating chemically synthesized restriction site linkers (adapters) [35] to cDNA was used in order to produce cDNA molecules with cohesive termini after digestion with a restriction endonuclease enzyme. Ligating the resulting cDNA to plasmid DNA cut with the identical restriction enzyme would create a recombinant plasmid which could then be cloned in a suitable bacterial host. Specifically, a decamer linker containing a site for HindIII was covalently joined to the ends of the cDNA with T4 DNA ligase, and then cleaved with HindIII; pMB9, a 3.5 million dalton plasmid conferring tetracycline resistance with a single site for HindIII, was also cut with the same endonuclease.

The usual procedure would be to now straightforwardly ligate these two molecules together creating the desired recombinant molecule. Kinetic theory [11] suggests that in order to insure ligation of most of the cDNA molecules to plasmid DNA, it is necessary to add a molar excess of plasmid DNA. However this would result in the majority of the plasmids simply self-circularizing without an insert of cDNA, and thus most the transformed cells would contain only pMB9 and not the desired recombinant plasmids. Here is where the novel step of removing terminal phosphates on the plasmid was generated.

Several sources of knowledge need be brought to bear in order to understand the basis of this new optimization step. Firstly, we need to know that alkaline phosphatase removes 5' terminal phosphates from the HindIII endonuclease-generated ends of the plasmid. Secondly, knowledge about the requirement of the T4 ligase for a phosphate end configuration allows us to infer that removing the phosphate ends prevents self-ligation of the plasmid DNA.

Thirdly, the kinetic theory of ligation [11] combined with a rule that says, "In a process that involves two or more competing components, you can optimize one process by inhibiting the other(s)", should tell us that circle formation is now dependent on the insertion of a DNA fragment containing 5'-phosphorylated termini: the cDNA. Finally, since transformation is directly linked to the DNA source, the one step inference is: "Only recombinant plasmids will transform the recipient bacteria".

A side effect that needs to be dealt with is the fact that the recombinant plasmids generated after phosphatase and ligase treatments will have two nicks, represented as asterisks in the figure, and that this has no known effect on transformation efficiency.

The application of alkaline phosphatase to remove terminal phosphates from a restriction enzyme-cleaved vector (e.g. plasmid) to eliminate self-ligation is a novelty that "should" have been obvious to any investigator working in this field. In fact, three or four years passed before Ullrich et. al. utilized it. One can only speculate as to the reasons why. However two related responses arise in this context. First, there are a very large number of DNA reagents available to the investigator (enzymes, chemical and separative techniques) so the number of possible combinations are vast. Secondly, especially with a well focused goal such as the ligation optimization step discussed above, people tend to think along relatively stereotyped paths, e.g. previously developed protocols. A computer system, such as MOLGEN, with a complex knowledge base and a good set of heuristic rules, will be likely to uncover novel applications of well known tools, precisely along the lines of the example just presented.

should tell us that circle formation is now dependent on the insertion of a DNA fragment containing 5'-phosphorylated termini: the cDNA. Finally, since transformation is directly linked to the DNA source, the one step inference is: "Only recombinant plasmids will transform the recipient bacteria".

A side effect that needs to be dealt with is the fact that the recombinant plasmids generated after phosphatase and ligase treatments will have two nicks, represented as asterisks in the figure, and that this has no known effect on transformation efficiency.

The application of alkaline phosphatase to remove terminal phosphates from a restriction enzyme-cleaved vector (e.g. plasmid) to eliminate self-ligation is a novelty that "should" have been obvious to any investigator working in this field. In fact, three or four years passed before Ullrich et. al. utilized it. One can only speculate as to the reasons why. However two related responses arise in this context. First, there are a very large number of DNA reagents available to the investigator (enzymes, chemical and separative techniques) so the number of possible combinations are vast. Secondly, especially with a well focused goal such as the ligation optimization step discussed above, people tend to think along relatively stereotyped paths, e.g. previously developed protocols. A computer system, such as MOLGEN, with a complex knowledge base and a good set of heuristic rules, will be likely to uncover novel applications of well known tools, precisely along the lines of the example just presented.

References

1. Bobrow D.G. and Winograd T., Experience with KRL-0: One Cycle of a Knowledge Representation Language, 5IJCAI:213-222 (1977)

2. Bobrow D.G. and Winograd T., An Overview of KRL, a Knowledge Representation Language, Cognitive Science 1 (1977)

3. Brachman R.J., What's in a Concept: Structural Foundations for Semantic Networks, BBN Report No. 3433 (1976)

4. Brown R., Use of Analogy To Achieve New Expertise, Artificial Intelligence Laboratory, MIT, AI-TR-403 (1977)

5. Buchanan, B.G. and Dennis Smith, "Computer Assisted Chemical Reasoning", in Proceedings of the III International Conference on Computers in Chemical Research, Education and Technology, Plenum Publishing, pp. 388-408 (1977)

6. Carhart, R.E., Smith, D., Brown, H., and Djeralls, C., Applications of artificial intelligence for chemical inference. XVII. An approach to computer-assisted elucidation of molecular structure J. Am. Chem. Soc. 97:5755-5762 (1975).

7. Davis R., Applications of Meta Level Knowledge to the Construction, Maintenance and Use of Large Knowledge Bases, Stanford Computer Science Department Report No. STAN-CS-76-552 (1977)

8. Davis R., Generalized Procedure Calling and Content-Directed Invocation, SIGPLAN Notices, 12:45-54 (1977)

9. Dershowitz N. and Manna Z., The Evolution of Programs: A System for Automatic Program Modification, Stanford University Artificial Intelligence Laboratory (1976)

10. Dijkstra E.W., Notes on Structured Programming, in Dahl,O., Dijkstra, E.W. and Hoare, C.A.R. in Structured Programming, Academic Press (1972)

11. Dugaiczyk, A., Boyer, H. and Goodman, H., Ligation of EcoRl Endonuclease- generated DNA Fragments into Linear and Circular Structures, J. Mol. Bio. 96:171-184 (1975)

12. Ehrlich S.D., Bursztyn-Pettegrew H., Stroynowski I. and Lederberg J., Expression of the thymidylate synthetase gene of the Bacillus subtilis bacteriophage Phi-3-T in Escherichia coli, Proceedings of the National Academy of Sciences USA, 73:4145-4149 (1976)

13. Evans T.G., A Program for the Solution of Geometric Analogy Intelligence Test Questions, in First Order Mathematical Logic, Margaris, A., Blaisdell Publishing Company, Waltham, Mass. (1967)

14. Feitelson J.S. and Stefik M.J., A Case Study of the Reasoning in a Genetics Experiment, Heuristic Programming Project Working Paper 77-18, Computer Science Department, Stanford University (1977)

15. Fikes R.E. and Nilsson N.J., Learning and Executing Generalized Robot Plans, Artificial Intelligence, 3:251-288 (1972)

16. Fikes R.E. and Nilsson N.J., STRIPS: A New Approach to Applications of Theorem Proving to Problem Solving, Artificial Intelligence 2:189-208 (1971)

17. Feigenbaum E.A., The Art of Artificial Intelligence: I. Themes and Case Studies of Knowledge Engineering, 5IJCAI:1014-1029 (1977)

18. Gerhart S.L., What Goes Down Should Also Come Up: Some Issues about Abstraction, 5th Texas Conference on Computing Systems (1976)

19. Gerhart S. L., Knowledge About Programs: A Model and Case Study, Proc. of Intl. Conf. on Reliable Software, Los Angeles (1975)

20. Goldberg A. and Kay A., SMALLTALK-72 Instruction Manual, Xerox Corporation (1976)

21. Goldstein I.P. and Roberts R.B., NUDGE, A Knowledge-based Scheduling Program, 5IJCAI:257-263 (1977)

22. Hayes-Roth F. and Lesser V.R., Focus of Attention in the Hearsay-II Speech Understanding System, 5IJCAI:27-35 (1977)

23. Kling R.E., Reasoning By Analogy With Applications To Heuristic Problem Solving: A Case Study, Stanford Computer Science Department Report No. CS-216 (1971)

24. Martin N., Friedland P., King J. and Stefik M., Knowledge Base Management for Experiment Planning in Molecular Genetics, 5IJCAI:882-887 (1977)

25. McDermott D., Vocabularies for Problem Solver State Descriptions, 5IJCAI:229-234 (1977)

26. Miller M.L. and Goldstein I.P., Structured Planning and Debugging, 5IJCAI:773-779 (1977)

27. Minsky M.A., A Framework for Representing Knowledge, in Winston P (ed) The Psychology of Computer Vision, New York: McGraw-Hill (1975)

28. Notani, N.K. and Setlow, J.K., Mechanism of Bacterial Tranformation and Transduction, in Progress in Nucleic Acid Research and Molecular Biology 14:39-100 (1974)

29. Platt J.R., Strong Inference, Science 146:347 (1964)

30. Rieger C. and Grinberg M., The Declarative Representation and Procedural Simulation of Causality in Physical Mechanisms, 5IJCAI:250-256 (1977)

31. Roberts, R.J. CRC Critical Reviews of Biochemistry 4 :123-164 (1976)

32. Sacerdoti E.D., The Nonlinear Nature of Plans, 4IJCAI,:206-214 (1975)

33. Sacerdoti E.D., Planning in a Hierachy of Abstraction Spaces, 3IJCAI:412-422 (1973)

34. Schaffner K., Logic of Discovery and Justification in Regulatory Genetics, Stud. Hist. Phil. Sci. 4:349-385 (1974)

35. Scheller R.H., Dickerson R.E., Boyer H.W., Riggs A.D. and Itakura K., Chemical Synthesis of Restriction Enzyme Recognition Sites Useful for Cloning, Science 196:177-180 (1977)

36. Stefik M.J. and Martin N., A Review of Knowledge Based Problem Solving as A Basis for a Genetics Experiment Designing System, Stanford Computer Science Department Report No. STAN-CS-77-596 (1977)

37. Sussman G.J., The Virtuous Nature of Bugs, Proceedings of the AISB Summer Conference (1974)

38. Ullrich A., Shine J., Chirgwin J., Pictet R., Tischer E., Rutter W.J., and Goodman H.M., Rat Insulin Genes: Construction of Plasmids Containing the Coding Sequences, Science 196:1313-1319 (1977)

39. Walker D.E., Paxton W.H., Grosz B.J., Hendrix G.G., Robinson A.E., Robinson J.J., Slocum J., Procedures for Integrating Knowledge in a Speech Understanding System, 5IJCAI:36-42 (1977)

40. Wipke W.T., Computer-Assisted Three Dimensional Synthetic Analysis, in Computer Representation and Manipulation of Chemical Information, Wipke W.T. et.al. (eds.) John Wiley (1974)

41. Woods B., What's in a Link?, in Bobrow and Collins (eds.) Representation and Understanding (1975)